

- [Contents](#) |
- [Author index](#) |
- [Subject index](#) |
- [Search](#) |
- [Home](#)

Proceedings of the Eighth International Conference on Conceptions of Library and Information Science, Copenhagen, Denmark, 19-22 August, 2013

Facets: ersatz, resource and tag

[Martin H. Frické](#)

The University of Arizona, SIRLS, 1515 E First St., Tucson, AZ
85719, USA

Abstract

Introduction. Faceted classification appears to be of utmost importance.

Ersatz facets, resource faceting and tag faceting. The distinctions are drawn between facets and ersatz facets, and between faceted resources and faceted tags.

Single tag resource faceting and multiple tag information object faceting The basic features are explored of single tag tagging of physical resources and multiple tag tagging of information entities.

Kaiser. Jerome Kaiser's work is identified as exemplifying a faceted theory but yet one that goes beyond ersatz faceting. **Aspects.** Considerations of 'aspects' are used to extend faceting yet further.

More general faceting. Ideas are offered on general faceting and its properties.

Conclusions. Faceted classification can be developed considerably further than it has been thus far.

CHANGE FONT

Introduction

Faceted classification appears to be of utmost importance. Many identify it as being the only true way to classify information objects ([Broughton 2006](#); [Classification Research Group 1955](#)), and it is a current hot topic of research in information science ([Broughton 2004, 2006](#); [Buchanan 1979](#); [Ranganathan 1962, 1967](#); [Wilson 2006](#); [Gnoli 2008](#); [Vickery 1960, 2008, 1966](#); [La Barre 2006, 2010](#); [Foskett 1996](#); [Foskett 2003](#); [Gardin 1965](#); [Hjørland 2012](#); [Perreault 1969a](#); [Beghtol 2008](#)). [The notion of information object in use in this paper encompasses the various kinds of IFLA F.R.B.R. Group 1 Entities ([IFLA 2009](#)).]

Faceting can be discussed in terms of syntax (i.e. words, terms, strings, etc.) or in terms of semantics (i.e. meanings, concepts, etc.) The present paper uses the familiar Triangle of Meaning ([Hjørland 2007](#); [Stock 2010](#); [Frické 2012](#)) to focus on concepts and the concept vertex, and, in turn, understands concepts as abstract objects ([Frické 2012](#)) (see also ([Hjørland 2009](#); [Szostak 2011](#))). The paper also uses some symbolic logic to identify those concepts ([Frické 2012](#)) (see also ([Gnoli 2006](#); [Stock 2010](#))).

Then, the basic idea or observation of faceting is that often the components of a synthesized compound classification concept are, or can be, categorized or faceted. For example, the topic ‘18th Century France’ is composed of a time period and a place. One component is of the category period and the other of the category place. There is a focus from a period facet, and a focus from a place facet. The faceting technique uses the fact that there are kinds or categories of concepts ([Austin 1984](#); [Foskett 1977](#); [Lambe 2007](#); [Morville and Rosenfeld 2006](#); [Willetts 1975](#); [Vickery 1960, 1966](#); [Cheti and Paradisi 2008](#); [Slavic 2008](#)).

Within faceted classification, Frické ([Frické 2011, 2012](#)) has drawn the distinction between facets and ersatz facets. What is this distinction and why is it important to draw it?

Ersatz facets, resource faceting and tag faceting

Vanda Broughton explains faceted classification by means of a sock example ([Broughton 2004, 2006](#)). In this, there are socks, which are long or short, and they are black or white, and they are made of wool or cotton, and so on. There is here an underlying domain (of socks), then properties or attributes that the entities in the domain may possess (being long, short etc.). The properties themselves belong to different categories—there are Length properties, there are Colour properties, and there are Materials properties. The categories of the properties are the facets (or faces) and the values or the properties in the facets are the foci. The Categories or facets themselves are orthogonal and this means that values of the foci in one are independent of values in another (so, for example, that some socks are long does not compel those socks to be black or compel them to be made of cotton). In contrast, the foci within a facet are dependent, in fact they are mutually exclusive (so, for example, if some socks are long that very fact prevents them from being short). From a technical and theoretical point of view it is probably left open within the model as to whether foci within a facet are, or have to be, exhaustive (i.e. whether they cover all the possibilities), but it is more convenient to assume that they are (and to ensure this a catch-all property or focus can be added (for example, ‘all-the-other-socks-which-are-neither-long-nor-short’)). So, what a single facet does, with its foci, is to partition the underlying domain into exclusive and exhaustive subsets. Each facet, on its own, does

this (and the facets will typically partition differently one from another). Conceptually, at least, the facets can be applied sequentially to give faceted search or faceted narrowing. So, for example, a User or Patron or Customer can be invited to choose among Length (of sock); what they are doing is selecting a particular subset of the socks as a whole; that subset (say the short socks) is itself partitioned by the other facets (for example, by Colour); and then the User can choose a Colour of sock (from among the short socks). Notice here, first, that everything is socks—the domain is socks, the subsets are socks, and the subsets of the subsets are socks—and, second, that order of narrowing, or application of the facets, or ‘filtering’, is of no consequence (for instance, short black socks and black short socks are one and the same subset of socks). These last two properties mean that a Boolean Faceted Search box could be constructed. Keywords such as ‘black’ and ‘short’ could be used as identifiers for the relevant properties, and a search input of ‘black AND short’ or ‘short AND black’ would behave as expected. (Of course, input boxes require from the User recall not recognition. So the User interface could be improved by inviting to User to make a choice, to ‘recognize’, among facet menus, and their foci instead of having to recall a suitable input phrase.)

What has just been described is ersatz faceting. It is very common, extremely common. The Web is where it flourishes. Many online department stores, museums, or similar, use it. All of us know it very well. Ersatz faceting can be made a little more elaborate than has been depicted here. For example, searches of foci might be arranged hierarchically by means of labels or meta-labels; the Colour choice menu might first offer ‘Pastel Colours’, as opposed to ‘Bold Colours’; in this configuration ‘Pastel’ would not itself be one of the foci, instead it would be a means of organizing some of the foci.

What this suggests is that a distinction be made between faceted classification and faceted search or retrieval. Colours, or Colour properties, like being black or being white, are foci and are applied directly to socks—they are part of the classification. But a notion like Pastel Colour offers choices to the information architect. It can be treated as a second order property, or meta-property, so that pink, for example, has the property of being a pastel Colour. In this guise, being a pastel Colour is not part of the faceted classification of socks; however, it can play a role in a faceted search among socks—it guides the User to make a narrowing choice among sets of foci prior to choosing a particular focus. Alternatively, ‘being pastel Coloured’ can be taken as a first order classifying property that some socks might have. Care is needed with this approach because, for example, being pastel Coloured is not independent of being pink (as pairs of foci would need to be). Likely what would need to be done is for the facet, for example, the Colour facet, to be organized into an Aristotelian Classification Hierarchy, then the foci would be the leaves of this tree, and more general properties (such as being pastel Coloured) would be interior nodes (which are not themselves foci).

This account of ersatz faceting gives us a very reasonable understanding of what is happening with websites, and similar, which give the User selection by progressive filtering via orthogonal properties. [c.f., for example, Endeca, Flamengo, and Apache Solr ([Hearst 2008](#); [Smiley and Pugh 2011](#); [Zelevinsky et al. 2008](#))]

Notice this about the logic of ersatz faceting. The white socks are some socks, the long socks are some socks, and the long white socks are the intersection of those two previous sets. That the facets are ‘intersective’ means both that the order of application of the modifying foci is irrelevant and that logically simple Boolean operations are adequate. One way of depicting the logical structure is to use ‘set builder’ notation from symbolic logic, in which case the proper concept or intension would be.

{x: Socks(x)&Long(x)&White(x)}

or, if the underlying domain of socks is understood contextually,

{x: Long(x)&White(x)}

To sum up some of the features of ersatz faceting of resources in the real world:

- there is one domain,
- selection by a single focus, or combination of foci from different facets, merely identifies a subset of the domain,
- intersective,
- selection of a single focus during a search automatically rules out any choice of other foci from the same facet, by exclusivity, so the other foci no longer need to be offered as choices in a search interface of the ongoing search,
- the order in which narrowing or filtering operations are carried out is inconsequential as to the final resulting subset (i.e. the operations permute or are symmetric),
- that the operations permute means that the syntax for a faceted description language is relatively open (e.g. either of 'long black socks' or 'black long socks' would be fine).
- the narrowing operations lend themselves to representation as Boolean search (in particular ANDing the operations in any order),
- for Boolean search
 - the ANDing operations need only AND foci across different facets (because choice of a focus within a facet implicitly and automatically ANDs that choice with NOT of all the other foci within a facet),
 - OR operations across facets is semantically sound (for example, 'long OR white' identifies a subset of socks),
 - OR operations within a facet is semantically sound (for example, 'wool OR cotton' identifies a subset of socks),
 - in fact, all the Boolean operations are sound, provided that the mutual exclusivity of foci within a facet is respected.

Single tag resource faceting and multiple tag information objects faceting

A table, possibly multi-dimensional or n-dimensional, is a very natural depiction of a faceted resource classification (see also ([Perreault 1969b](#))). For example, say the socks had two facets: Length, and Colour; one of these facets could be the columns in a table and other facet the rows, then each cell would be a combination of a row and a column, a combination of length and Colour.

Table 1: A faceted resource classification

	long	medium	short
black			
pink			
white			

Each of the cells is a most specific kind that each individual pair of socks might have. So if there were identifiers for the individual pairs of socks, say #1, #2, #3, #4, #5. Their classification might be depicted as

Table 2: Some classified resources

	long	medium	short

black	5		3
pink		1,2	
white	4		

This generalizes to the many dimension case with more than two facets. In this vein, any cell can be picked out by a vector, which is just an ordered combination of the foci, for example

long, white, cotton

might pick a specific narrowest kind or subset or cell of socks. [There is a mild disanalogy in the use of tables or vectors to portray facets. Tables, and vectors, have order (e.g. rows then columns) but facets need not (e.g. it does not matter whether the Length of socks is depicted in the rows or, alternatively, in the columns).]

With search, for example in a two dimensional table, narrowing could be selecting a single column (say ‘long’, which is a single focus) or by selecting a single row (say ‘white’, which is also a single focus). But it also could be by selecting several rows at once, or several columns at once. That, for example, is what is happening with a search choice like ‘Pastel Colours’ (which is not a focus at all, rather it is something of a Boolean disjunct or combination of rows or foci). Again, this generalizes to multiple facets (multi-dimensional tables, or vectors of any length).

With the socks example, it is a resource (namely the socks) that is being faceted. But there is a different way we can think about this. We can imagine that each pair of socks carries a single tag which identifies what kind of pair of socks it is; here are some sample tags

long, white, cotton

long, white, synthetic

long, white, wool

long, black, cotton

long, black, synthetic

long, black, wool

short, white, cotton

etc.

These tags, collectively, form a tagging language. And each pair of socks is labeled with exactly one of them (which informs of its most specific kind). The tagging language itself is, or can be conceived to be, faceted—the tags are permutations of exclusive foci from orthogonal facets. (Travis Wilson makes the distinction between resource faceting and tag faceting (and makes the likely correct observation that Ranganathan’s work was with tag faceting ([Wilson 2006](#))) See also ([Gnoli 2006](#); [Svenonius 2000](#)).)

As Information Scientists, we are not particularly interested in socks. What we are interested in is books about socks (and other topics), scholarly articles about socks, DVDs about socks, and so forth. In short, our professional interest is with information objects about socks (and other topics). What

does ersatz faceting have to offer us by way of helping to classify or retrieve those? Certainly we can use our real world ersatz scheme to produce for us labels or 'tags' or 'subject headings' and then use those to tag the relevant information objects. So, for example, 'socks', 'black socks', 'long socks', 'long black socks', and so on could be tags which are used to tag the information objects which address those topics. This scheme could also support faceted search or faceted retrieval.

Notice here the double duty that a tag language might play. The tag

long, white, cotton

can be used to identify a kind for a pair of socks; but the same tag (and its associated entire tag language) can be used homographically, or punningly, to identify a topic (or topics) that an information object addressed. (Elaine Svenonius calls this the 'referential semantics' of subject languages (([Svenonius 2000](#))p.130).)

There is a point to be made about the grammar of ersatz faceted tags. Subject tags grammatically are going to be nouns, and compound tags typically are nouns modified by 'modifiers' such as adjectives, noun phrases, etc. ; so, for example, 'long white socks' is the noun 'socks' modified into a compound noun by the two modifying adjectives 'long' and 'white'. That the underlying ersatz facets are 'intersective' accompanies the fact that the modifiers are intersective.

An entirely separate, but nonetheless important, question is whether an information object has to carry exactly one tag (as the socks do) or whether it is permitted to carry many tags. Why this matters is whether we plan to shelve the information objects using the tag scheme. If we plan to shelve the information objects, then there has to be a single or primary subject or tag for the information object to generate a unique Call Number to shelve the information object. This leads us into considerations of citation order (for example, do we prefer the syntax 'black socks' to 'socks black') and then to filing order among the foci of each facet (for instance, should 'black' come before 'white'). Many theoreticians have gone deep into this; Ranganathan, for one, did. But it faces an unsurmountable problem. Many or most information objects are on two or more topics; and forcing them to appear to be on one topic is totally artificial and is carried out solely for the purposes of shelving. It follows that shelving, in the sense of perfection of location and collocation simply cannot be done. As Jesse Shera observed

...the history of library classification...has been the narrative of a pursuit of impossible goals. ([Shera 1965](#))p.100

A conclusion is: do not do it, do not aim at single, or primary, tag tagging. Almost all information object tagging should be multiple-tags-possible tagging; that is, it should be what librarians call 'subject classification'.

With subject tagging, information objects can carry multiple tags. In particular, an information object can carry what appear to be contradictory tags, for example 'black socks', and 'white socks'. This merely means that the information object in question addresses two or more topics whose real world entity designations are distinct and mutually exclusive. How to design an ersatz faceted topic tag scheme is a bit of an open question, especially as to its details. But, broad brush, one approach would be to use something similar to Web site faceting of objects but to permit multiple tags.

A similar table could be used, but now information objects, for example, #a, #b, #c, #d, #e might be in more than one cell

Table 3: Some tagged information objects

	long	medium	short
black	e		c
pink		a,b	d
white	e		e

And this might signify that the information object ‘a’ address the single topic

medium, pink

and the information object, ‘e’ addressed the three topics

long, black

long, white

short, white

Yet again this generalizes easily to any number of facets.

Search or retrieval might be a little bit more elaborate than previously but it would not be different in principle. [For example, care would have to be taken over the syntax and semantics of a Boolean Search Box; for instance a retrieval of ‘[topic] black socks AND white socks’ presumably would be understood to mean or designate the subset of those information objects which carry both the ‘black socks’ tag and the ‘white socks’ tag, whereas a retrieval of ‘[topic] black and white socks’ would fail, or retrieve nothing, because the tag ‘black and white socks’ is ill-formed, there is no such proper tag.

There is still the characteristic narrowing of ersatz faceting. If a search for the topic long produces a set of information objects on the topic of long socks, then if the topic is narrowed to long white that set of information objects will be narrowed to one of its subsets.

To sum up some of the features of multiple tag tagging using ersatz faceted subject tags:

- the topics are about one domain,
- selection by a single focus, or combination of foci (including sums of combinations of foci) from the same or different facets, merely identifies a subset of the information objects,
- intersective,
- selection of a single focus during a search does not rule out any choice of other foci from the same facet, so the other foci do need to be offered as choices in a search interface for the ongoing search,
- the order in which filtering operations are carried out is inconsequential as to the final resulting subset (i.e. the operations permute or are symmetric),
- the narrowing operations lend themselves to full representation as Boolean searches (using AND, OR, and NOT operations across foci from any facet whatsoever).

That two or more foci from within the same facet can be combined within a search (for example, the [tag and topic] white socks together with the [tag and topic] black socks) opens up the possibility for a different design of tag language in which the combinations occur actually within a single tag. For example, consider a book on 17th and 18th Century France; with a basic type of tag language that would need to carry two tags, one in effect ‘17th Century France’ and the other ‘18th Century

France'; but a more elegant design might just have the single tag, in effect '17th and 18th Century France'. And certainly some theoreticians acknowledge and accommodate this, for example Barbara Kwasnik writes

Each facet can be developed/expanded using its own logic and warrant and its own classificatory structure. For example, the Period facet can be developed as a timeline; the Materials facet can be a hierarchy; the Place facet a part/whole tree, and so on. ([Kwasnik 1999](#))p.39-40

Kaiser

Julius Otto Kaiser was one of the early figures to use faceting ([Kaiser 1911](#); [Dousa 2011](#)). He used it in indexing, which we take as being essentially the same as subject classification. Kaiser's area of employment was in commercial and industrial libraries, in particular those relating to heavy industry. He noticed that many of the subject topics of interest could be composed out of instances of three kinds of categories: Concretes, Processes, and Countries. Concretes were 'things', general or specific (i.e. types or tokens ([Wetzel 2009](#))) such as aluminum, iron, and steel; and Processes included smelting, refining, and rusting; and Countries were locations (and we need not consider Countries further in this paper). So a subject tag might be 'Aluminum refining'. This tag identifies a compound concept and it is made out of a value or focus from the Concretes facet together with a focus from the Processes facet.

But what is going on here is somewhat different to ersatz faceting. Let us start in the real world of actual things or resources (prior to worrying about subjects or topics). What might be the domain of a Kaiser scheme? Suppose it is the Concretes i.e. a collection of things like aluminum, iron, steel, etc. What is the effect of combining Concretes with a Process, say the focus 'refining'? It is not to create a subset of the Concretes. 'Aluminum refining', for example, is not a Concrete at all (in fact, it is a Process). The faceted combination leaves the original domain. Combining foci from different facets is not narrowing or filtering. Kaiser is creating a synthetic subject tagging or indexing language. That the subject language is synthetic requires that it have a syntax or grammar, and Kaiser produces one in terms of combinations of foci from facets. But the constructions are not filtering constructions.

Thomas Dousa tells us that Kaiser, when indexing, would have in effect used processes as subterms for concretes (([Dousa 2011](#))p.167). And he quotes Kaiser as writing

To put it into the simplest language we may say that literature names things and that these things are spoken of or described. The knowledge conveyed by literature all has reference either to *things* or to *spoken of*, i.e. concretes and processes (§ 298 [emphases his]).
(([Dousa 2011](#))p.169)

And then remarks that the Kaiser view here might be identified as that of ascribing a property or predicate (the process) to a subject (the concrete) (([Dousa 2011](#))p.169). Those views—that process terms are properly subterms of concrete terms and combining a concrete with a process is just applying a predicate to the concrete—are just mistaken (whether or not, as a matter of exegesis, Kaiser, or anyone else, actually held them, or, indeed, if Dousa himself subscribes to them).

To depict this using set builder notation, if the iron Concretes are

$$\{x: \text{Iron}(x)\}$$

then the smelting iron Process *cannot* be

$$\{x: \text{Iron}(x) \& \text{Process}(x)\}$$

because the Iron(x) property picks out iron ingots not processes. Instead the concept smelting Iron has to be something like

$$\{x: \text{Process}(x) \& (\exists y(\text{Iron}(y) \& \text{Smelts}(x,y)))\} \text{ (i.e the processes such that there are iron ingots which they smelt)}$$

But we could have taken the original domain to a sorted domain that had the two sorts, Concretes and Processes, in it. But really this does not help. The Concretes are a subset of the two-sorted domain; when a Process focus is applied to it, that application results in a subset of the Process subset of the two-sorted domain; it does not ‘narrow’ on the original Concretes subset. This is not like the narrowing successive subset search of ersatz faceting. The faceting is not intersective.

We might alternatively say, or have said, that the Processes are the domain. So all there is in the world is refining, smelting, rusting, etc. Then Aluminum refining would also be Process, a narrowed subset of the original domain. That would work to a degree. But the combinations would be ‘subsective’ (not intersective). For example, when iron is combined with smelting what results is a subset or subset of smelting only, a Process, and not simultaneously a subset of the Concretes. [Subsecting on Processes is also just wrong as to what is being attempted. Kaiser would have said that a Concrete on its own (say iron) would be a perfectly good topic, so, in the world, Concretes need to be in the domain as first class citizens in their own right and not just as components of processes.]

We can see what the challenge is by attempting to show it in an ersatz-faceted style of table, say

Table 4: An attempt at a classification table

	aluminum	copper	iron
anneal			
refine			
smelt			

As depicted here, all the cells are processes. In which case, what has happened to the plain Concretes like iron? We could have a second table for them, in which case we have moved on to a two-sorted domain.

Kaiser faceting cannot be used to organize a website in the same simple way as ersatz faceting. Concretes and Processes are good deal more abstract than, for example, socks. But if we suspend our disbelief for one moment, and let our imagination run. We could imagine a Foundry that sells items or services to the public via a website. The website might have a Concretes menu with foci aluminum, iron, etc., and in conjunction with that menu there were other menus, or facets, such as Price-per-ingot with foci \$100, \$100-\$500, etc, and Availability with foci December 2013, January 2014 etc. That would be ersatz faceting, with its familiar narrowing of subsets. But suppose the website also had a Processes menu and facet with foci like smelting, annealing, etc. which were services that the Foundry offered. If this were used in conjunction with the Concretes menu, it would not be to narrow on subsets of ingots. It might be conceived as narrowing on Processes. But then the website as a whole would be ersatz faceted on two distinct domains: Concretes and Processes. There would be four facet menus: Concretes, Price-per-ingot, Availability, and Processes. The first three could be used sequentially in any order to narrow on sets of ingots (they operate on the sort of Concretes). The first

and the last could be used presumably in some conventional order to narrow on Processes (the work on the sort of Processes). The last three—Price-per-ingot, Availability, and Processes—presumably cannot be used together, at least not without change of their original semantics. And Boolean Search, were it to be attempted, comes under stress because, for example, ‘smelting iron’ simply is not ‘smelting AND iron’.

Under this analysis of resources in the world,

- there might be multiple domains, or domains containing multiple sorts,
- some operations on these domains or sorts might pairwise be intersective, but some others also might individually be subsective,
- selection by a single focus, or combination of foci, might identify a subset of one of the domains, but it might transfer from one domain to another,
- the order in which ‘narrowing’ or filtering operations are carried out might matter (i.e. the operations might not permute or might not be symmetric or might not be well formed),
- there may be difficulties in depicting the search process as a simple Boolean Search.

The Kaiser scheme, were it to be used for subject tagging, would represent something of a generalization of ersatz tagging. But there is a motivation not to rest here but rather to generalize yet further, and that motivation is provided by ‘aspects’.

Aspects

It is regularly observed that topics for the purposes of subject classification often involve ‘aspects’. Brian Vickery writes

Taxonomy is basically concerned with classifying ‘natural kinds’—of organisms, of soils, of substances. Documentation has to classify what is written about these objects, and must take into account not only the natural kinds but also their properties, behaviour, interactions, and operations on them. (([Vickery 1975](#))p.9)

The point here is that a topic might be the simple atomic concept, say iron, but associated with this are a number of compound concepts such as, the smelting of iron, the refining of iron, the mining of iron, the use of iron in construction, the value of iron to an economy, the use of iron in art, iron in tools, and so on. These compound concepts are ‘aspects’, in this case ‘aspects of iron’. And, typically, it is ‘aspects’ that are important to the information scientist. [Explanations of ‘aspects’ can be found in ([Broughton 2004](#); [Broughton et al. 2005](#); [Hjørland 2006](#); [Mills and Broughton 1977](#)).]

Topics, as aspects, are going to be compound concepts or compound nouns, some of which, as noted, can be conceived as being intersective or subsective and these can be used for narrowing faceted search. But there are also aspect combinations of nouns and modifiers which do not seem to have these properties (for example, a ‘former student’ is not a student so there is no starting with students and narrowing to former students (although there is the workaround of starting with past, present, and future students and narrowing to former students)). Linguists talk of non-predicating modifiers, and these simply do not seem to be either intersective or subsective.

A not untypical example of an aspect provided by Broughton is: ‘rabbits in farming’ ([Broughton 2004](#)). It is hard to see this as being intersective or subsective of any simple properties of anything, rather it is a compound something like

$\{x: (\exists y(\text{Rabbit}(y)) \ \& \ \exists z(\text{Farming}(z)) \ \& \ \text{Use}(x,y,z)) \}$ (i.e the uses such that they are uses of rabbits in farming)

Even so, a compound like this could be constructed from faceted components—the rabbits, farming, and uses could all be foci of more general categories.

More general faceting

So, in the general case, there can or might be suitable classificatory compound concepts which have faceted construction. And this means that they are made out of components, eventually atomic components, which have different kinds or are of different categories.

The general case includes ersatz faceting but goes beyond it. The most obvious example of going beyond it is where there are functions involved and not just simple properties. Consider the following as a classification of real world resources. First, Monarchs of England, this is a perfectly good set of people, namely a collection of Kings and Queens. But now move on to Spouses of Monarchs of England, this also is a perfectly good set of people, but these are different people, not tending to be Monarchs (and then Relatives of Spouses of Monarchs are yet further people still). When someone is a spouse, say Phillip, the present so-called Prince Phillip, he certainly has the property of being a spouse, but that plain property does not capture all the information for Phillip is not merely a spouse he is the spouse of Elizabeth II (and that invites the use of a function for ‘is a spouse of’). A simple faceted classification could capture this:

Facet 1: Monarchs (with foci Charles, Elizabeth, George, Henry etc.)

Facet 2: Direct family relationship (with foci spouse, mother, father etc.)

Facet 3: (Less) Direct family relationship (with foci spouse, mother, father etc.)

So, for example, Monarchs, Monarch’s spouses, Monarch’s spouse’s mothers, Monarch’s mother’s spouses are all perfectly good classification categories whose contents will almost be different one from another, indeed quite different.

Exactly the same consideration applies with information objects and information object tagging using functions. If, for example, we used a faceted classification to retrieve all the information objects tagged ‘Henry VIII’ and then decide we want instead information objects on the topic of ‘the wives of Henry VIII’, that second collection of information objects, which we get from using the constructed faceted tag ‘Henry VIII spouses’ will not be a subset of the first. There is not intersection or subsection going on here. There is no narrowing or filtering, and the order of application of the various foci matters (a mother’s spouse and a spouse’s mother would usually be different). This faceting is entirely different to ersatz faceting and much stronger. And it is likely also stronger than ‘aspects’: the wives of Henry VIII seem hardly to be an aspect of Henry VIII.

Conclusion

Ersatz faceting typically uses just AND in set builders, for example,

$\{x: \text{Long}(x) \ \& \ \text{White}(x)\}$

That is fine, and it is especially good for organizing websites of things (see also ([Gnoli 2006](#))). But for information science faceted classification can and should be taken beyond this—to combinations of atomic concepts that use full predicate calculus operations including quantifiers and functions.

If we look back to the history of faceting in librarianship and information retrieval, some work of particular interest includes that by Derek Austin, Jason Farradane, J-C Gardin, and J.M. Perreault ([Austin 1984](#); [Farradane 1950, 1963](#); [Gardin 1965, 1969, 1973](#); [Perreault 1965, 1969a](#); [Moss 1964](#); [Campbell 2011](#)). Essentially they faceted down to foci but then introduced the additional step of having (simple) relations between foci. So that, for example, ‘explosions’ might be one focus, ‘injuries’ another, and ‘...causing...’ a relation between certain kinds of foci, and these components could be put together to form the subject tag ‘explosions causing injuries’. This is admirable work, but from a logical point of view it could be taken a lot further. For example, there could be quantification over the causes or effects, say ‘anything causing injuries’. And the relations themselves could be categorized and possibly faceted.

What is needed here, and one way to generalize, is for there to be a logically detailed description of the world, an ontology. And, in fact, there are such, or attempts at such. There is SUMO, BFO, DOLCE, and other general ontologies ([Grenon 2003](#); [Masolo et al. 2001](#); [Niles and Adam 2001](#); [Khoo and Na 2006](#)). These are not subject tagging languages. Primarily they are to identify what is in the world so that information about those items can be entered into databases and so that the resulting databases can communicate one with another perhaps as part of the semantic web. But the ontologies could be understood as subject tagging languages via the referential semantics device mentioned earlier. And likely they could be faceted. What these general ontologies do not do, or do not do a lot, is to address aspects. They are principally interested in things, and certainly interested in some, or even many, properties of things, but they do not focus on aspects. So, a way forward with subject tagging might be to take some of the general ontologies and to add faceting and aspects to them. In sum, work remains to be done.

About the author

Martin Frické is an Associate Professor in the School of Information Resources and Library Science, The University of Arizona, USA. He received his Bachelor's degree in Philosophy from the University of Exeter, UK, his Bachelor's degree in Computer Science from the University of Otago, New Zealand, his Masters and his PhD degrees from the London School of Economics, UK. He can be contacted at: mfricke@u.arizona.edu.

References

- Austin, D. (1984). *PRECIS: a manual of concept analysis and subject indexing* (2 ed.). London: The British Library
- Beghtol, C. (2008). From the Universe of Knowledge to the Universe of Concepts: The Structural Revolution in Classification for Information Retrieval. *Axiomathes*, **18**(2), 131-144
- Broughton, V. (2004). *Essential Classification*. New York: Neal-Schuman
- Broughton, V. (2006). The need for a faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings: New Information Perspectives*, **58**(1/2), 49-72
- Broughton, V., Hansson, J., Hjørland, B., and López-Huertas, M. J. (2005). Knowledge organization. In L. Kajberg and L. Lørring (Eds.), *European curriculum reflections on library*

- and information science education (pp. 133-148). Copenhagen: Royal School of Library and Information Science
- Buchanan, B. (1979). Theory of Library Classification. *London: Clive Bingley*
 - Campbell, D.G. (2011). Revisiting Farradane's Relational Indexing in a Consumer Health Context. Retrieved 15 July, 2013 from <http://www.iskouk.org/conf2011/> (Archived by WebCite® at <http://www.webcitation.org/6IeUM9HqJ>)
 - Cheti, A., & Paradisi, F. (2008). Facet analysis in the development of a general controlled vocabulary. *Axiomathes*, **18**(2), 223-241
 - Classification Research Group. (1955). The need for a faceted classification as the basis of all methods for information retrieval. *Library Association Record*, **57**(7), 262-268
 - Dousa, T.M. (2011). Concretes, countries, and processes in Julius O. Kaiser's Theory of Systematic Indexing: A case study in the definition of general categories. In R. P. Smiraglia (Ed.), *Proceedings from North American Symposium on Knowledge Organization* (Vol. 3, pp. 160-173). Toronto, Canada
 - Farradane, J.E.L. (1950). A scientific theory of classification and indexing. *Journal of Documentation*, **6**, 83-99
 - Farradane, J.E.L. (1963). Relational indexing and new methods of concept organization for information retrieval *American Documentation Institute 26th Annual Meeting* (pp. 135-136)
 - Foskett, A. C. (1977). Subject approach to information. (3 ed.). *London: Clive Bingley*
 - Foskett, A. C. (1996). Subject approach to information (5 ed.). *London: Facet Publishing*
 - Foskett, D. J. (2003). Facet analysis. In M. A. Drake (Ed.), *Encyclopedia of Library and Information Science* (2 ed., pp. 1063-1067). New York: Marcel Dekker
 - Frické, M. (2011). Faceted classification: orthogonal facets and graphs of foci? *Knowledge Organization*, **38**(6), 491-502
 - Frické, M. (2012). Logic and the Organization of Knowledge. *New York: Springer*
 - Gardin, J.-C. (1965). Free classifications and faceted classifications: their exploitation with computers (J. E. L. Farradane, Trans.). In P. Atherton (Ed.), *Classification research: proc. int. conf. Elsinore 1964*, Munksgaard, Copenhagen, 1965 (pp. 161-176)
 - Gardin, J.-C. (1969). Semantic analysis procedures in the sciences of man. *Social Science Information*, **8**(17), 17-42
 - Gardin, J.-C. (1973). Document analysis and linguistic theory *Journal of Documentation*, **29**(2), 137-168
 - Gnoli, C. (2006). The meaning of facets in nondisciplinary classifications. Paper presented at the 9th ISKO conf. Vienna
 - Gnoli, C. (2008). Facets: A Fruitful Notion in Many Domains. *Axiomathes*, **18**(2), 127-130
 - Grenon, P. (2003). BFO in a Nutshell: A Bi-categorical Axiomatization for BFO and Comparison with DOLCE: IFOMIS Report 06/2003 December 2003
 - Hearst, M. (2008). UIs for Faceted Navigation: Recent Advances and Remaining Open Problems. *Workshop on Human-Computer Interaction and Information Retrieval, HCIR 2008*, Redmond, WA
 - Hjørland, B. (2006). Aspect or discipline versus entity or phenomena or "one place" classification. Retrieved July 15 2013 from http://www.iva.dk/bh/lifeboat_ko/concepts/aspect_classification.htm (Archived by WebCite® at <http://www.webcitation.org/6IeUy2R8q>)
 - Hjørland, B. (2007). Semantics and knowledge organization *Annual Review of Information Science and Technology* **41**, pp. 367-405. Medford, N.J.: Information Today, Inc.
 - Hjørland, B. (2009). Concept theory. *Journal of the American Society for Information Science and Technology*, **60**(8), 1519-1536.

- Hjørland, B. (2012). Facet analysis: The logical approach to knowledge organization. *Information Processing and Management*
- International Federation of Library Associations and Institutions (2009) *Functional Requirements for Bibliographical Records*
http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf
- Kaiser, J. O. (1911). Systematic Indexing. *London: Pitman*
- Khoo, C.S.G., and Na, J.C. (2006). Semantic relations in information science. *Annual Review of Information Science and Technology*, **40**(1), 157-228
- Kwasnik, B.H. (1999). The Role of Classification in Knowledge Representation and Discovery. *Library Trends*, 48(1), 22-47
- La Barre, K. (2006). The use of faceted analytico-synthetic theory as revealed in the practice of website construction and design. *Indiana University, Bloomington*
- La Barre, K. (2010). Facet Analysis. In B. Cronin (Ed.), *Annual Review of Information Science and Technology* (pp. 243-286). Medford, NJ: Information Today, Inc.
- Lambe, P. (2007). Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness. *Oxford, England: Chandos Publishing*
- Masolo, C., Borgo, S., Gangemi, A., Guarino, N., and Oltramari, A. (2001). WonderWeb Deliverable D18. from <http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf>
- Mills, J., and Broughton, V. (1977). Bliss Bibliographic Classification. Second Edition. Introduction and Auxiliary Schedules. *London: Butterworths*
- Morville, P., and Rosenfeld, L. (2006). Information Architecture for the World Wide Web. *Sebastapol, CA: O'Reilly*
- Moss R (1964) Categories and relations: Origins of two classification theories. *American Documentation* **15** (4):296-301
- Niles, I., and Adam, P. (2001). Towards a Standard Upper Ontology. Paper presented at the FOIS '01 *Proceedings of the international conference on Formal Ontology in Information Systems* pp 2-9, Ogunquit, Maine, USA. ACM New York
- Perreault, J.M. (1965). Categories and relators: a new schema [Reprinted in his book]. *Revue internationale de documentation*
- Perreault, J.M. (1969). *Towards a theory for UDC*. London: Archon Books & Clive Bingley
- Perreault, J.M. (1969). Transparency and self-definition in classification. In J.M. Perreault (Ed.), *Towards a theory for UDC* (pp. 90-118). London: Archon Books & Clive Bingley
- Ranganathan, S.R. (1962). *Elements of library classification* (3 ed.). Bombay: Asia Publishing House
- Ranganathan, S.R. (1967). Prolegomena to Library classification. 3. Retrieved July 15, 2013 from <http://dlist.sir.arizona.edu> (Archived by WebCite® at <http://www.webcitation.org/6leVSfXXI>)
- Shera, J.H. (1965). Classification: Current functions and applications to the subject analysis of library materials. *Libraries and the organization of knowledge* (pp. 97-111). Hamden, CT: Archon
- Slavic, A. (2008). Faceted Classification: Management and Use. *Axiomathes*, **18**(2), 257-271
- Smiley, D., and Pugh, E. (2011). Apache Solr 3 Enterprise Search Server (1 ed.): *Packt Publishing*
- Stock, W.G. (2010). Concepts and semantic relations in information science. *Journal of the American Society for Information Science and Technology*, **61**(10), 1951-1969
- Svenonius, E. (2000). *The Intellectual Foundation of Information Organization*. Cambridge, MA: M.I.T. Press
- Szostak, R. (2011). Complex concepts into basic concepts. *Journal of the American Society for Information Science and Technology*, **62**(11), 2247-2265

- Vickery, B.C. (1960). *Faceted classification: a guide to construction and use of special schemes*. London: Aslib
- Vickery, B.C. (1966). Faceted classification schemes. In S. Artandi (Ed.), *Rutgers Series on Systems for the Intellectual Organization of Information* (Vol. 5). New Brunswick, NJ: Graduate School of Library Science at Rutgers University
- Vickery, B.C. (1975). Classification and indexing in science (3 ed.). *London: Butterworths*
- Vickery, B. C. (2008). Faceted Classification for the Web. *Axiomathes*, **18**(2), 145-160
- Wetzel, L. (2009). Types and Tokens: On Abstract Objects. Cambridge, MA: *MIT Press*
- Willetts, M. (1975). An investigation of the nature of the relation between terms in thesauri. *Journal of Documentation*, **31**(3), 158-184
- Wilson, T. (2006). The strict faceted classification model. Retrieved 15 July, 2013 from <http://facetmap.com/pub/> (Archived by WebCite® at <http://www.webcitation.org/6IeVk41qL>)
- Zelevinsky, V., Wang, J., and Tunkelang, D. (2008). Supporting Exploratory Search for the ACM *Digital Library Workshop on Human-Computer Interaction and Information Retrieval* (HCIR'08)

How to cite this paper

Frické, M.H. (2013). Facets: ersatz, resource and tag. *Information Research*, **18**(3) paper C25. [Available at <http://InformationR.net/ir/18-3/colis/paperC25.html>]

Find other papers on this subject

[Scholar Search](#)
[Google Search](#)
[Bing](#)

Check for citations, [using Google Scholar](#)

[Like](#)

0

[Tweet](#)

7

223

© the author, 2013.

Last updated: 11 August, 2013

- [Contents](#) |
- [Author index](#) |
- [Subject index](#) |
- [Search](#) |
- [Home](#)